



-25-

# SEQUENCE LISTING

## GENERAL INFORMATION

- (i) APPLICANTS: NANCY CARRASCO, ET AL.
  - (ii) TITLE OF INVENTION: THYROID SODIUM/IODIDE SYMPORTER AND NUCLEIC ACID ENCODING SAME
  - (iii) NUMBER OF SEQUENCES: 2
  - (iv) CORRESPONDENCE ADDRESS:
    - (A) ADDRESSEE: AMSTER, ROTHSTEIN & EBENSTEIN
    - (B) STREET: 90 PARK AVENUE
    - (C) CITY: NEW YORK
    - (D) STATE: NEW YORK
    - (E) COUNTRY: U.S.A.
    - (F) ZIP: 10016
  - (v) COMPUTER READABLE FORM:
    - (A) MEDIUM TYPE: 3.5 INCH 1.44 Mb STORAGE DISKETTE
    - (B) COMPUTER: IBM PC COMPATIBLE
    - (C) OPERATING SYSTEM: MS-DOS
    - (D) SOFTWARE: ASCII
  - (vi) CURRENT APPLICATION DATA:
    - (A) APPLICATION NUMBER: 08/595,553
    - (B) FILING DATE: FEBRUARY 1, 1996
  - (viii) ATTORNEY/AGENT INFORMATION:
    - (A) NAME: CRAIG J. ARNOLD
    - (B) REGISTRATION NUMBER: 34,287
    - (C) REFERENCE/DOCKET NUMBER: 96700/393
  - (ix) TELECOMMUNICATION INFORMATION:
    - (A) TELEPHONE: (212) 697-5995
    - (B) TELEFAX: (212) 286-0854 or 286-0082
    - (C) TELEX: TWX 710-581-4766
- (2) INFORMATION FOR SEQ ID NO: 1
- (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 2839
    - (B) TYPE: NUCLEIC ACID
    - (C) STRANDEDNESS: DOUBLE
    - (D) TOPOLOGY: LINEAR
  - (ii) MOLECULE TYPE:
    - (A) DESCRIPTION: OLIGONUCLEOTIDE

COPY OF PAPERS  
ORIGINALLY FILED

- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: RAT  
 (B) INDIVIDUAL ISOLATE: SODIUM/IODIDE SYMPORTER
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1

GAATTCCGGG TCGACCACGC GTCCGGCGGT GACTCGCGCT GCGACTCTCC	50
CACTGACCGA GAGTCCCCGA CGTCCTCCGC ATCCTCTCCT CACCGAGTCA	100
CCTGTCTCC ATG GAG GGT GCG GAG GCC GGG GCC CGG GCC ACC	142
Met Glu Gly Ala Glu Ala Gly Ala Arg Ala Thr	
1 5 10	
TTC GGC GCC TGG GAC TAC GGC GTG TTC GCG ACC ATG CTG CTG	184
Phe Gly Ala Trp Asp Tyr Gly Val Phe Ala Thr Met Leu Leu	
15 20 25	
GTG TCC ACG GGC ATC GGG CTA TGG GTC GGC CTG GCC CGC GGT	226
Val Ser Thr Gly Ile Gly Leu Trp Val Gly Leu Ala Arg Gly	
30 35	
GGC CAA CGC AGT GCC GAC GAC TTC TTT ACC GGG GGC CGG CAG	268
Gly Gln Arg Ser Ala Asp Asp Phe Phe Thr Gly Gly Arg Gln	
40 45 50	
TTG GCA GCC GTT CCT GTG GGG CTG TCG CTG GCC GCC AGT TTC	310
Leu Ala Ala Val Pro Val Gly Leu Ser Leu Ala Ala Ser Phe	
55 60 65	
ATG TCG GCT GTG CAG GTG CTC GGG GTC CCC GCC GAG GCA GCG	352
Met Ser Ala Val Gln Val Leu Gly Val Pro Ala Glu Ala Ala	
70 75 80	
CGC TAC GGG CTC AAG TTC CTG TGG ATG TGC GCG GGT CAG TTG	394
Arg Tyr Gly Leu Lys Phe Leu Trp Met Cys Ala Gly Gln Leu	
85 90 95	
CTC AAC TCG CTG CTC ACA GCG TTT CTC TTC TTG CCG ATC TTC	436
Leu Asn Ser Leu Leu Thr Ala Phe Leu Phe Leu Pro Ile Phe	
100 105	
TAC CGC CTG GGC CTT ACC AGC ACC TAC CAG TAC CTA GAG CTG	478
Tyr Arg Leu Gly Leu Thr Ser Thr Tyr Gln Tyr Leu Glu Leu	
110 115 120	

CGC	TTC	AGC	CGA	GCG	GTC	CGG	CTC	TGC	GGG	ACG	CTG	CAG	TAC	520
Arg	Phe	Ser	Arg	Ala	Val	Arg	Leu	Cys	Gly	Thr	Leu	Gln	Tyr	
	125					130					135			
TTG	GTG	GCC	ACG	ATG	CTG	TAT	ACA	GGC	ATC	GTG	ATC	TAC	GCG	562
Leu	Val	Ala	Thr	Met	Leu	Tyr	Thr	Gly	Ile	Val	Ile	Tyr	Ala	
		140					145					150		
CCT	GCG	CTC	ATC	CTG	AAC	CAA	GTG	ACC	GGG	TTG	GAC	ATC	TGG	604
Pro	Ala	Leu	Ile	Leu	Asn	Gln	Val	Thr	Gly	Leu	Asp	Ile	Trp	
			155					160					165	
GCA	TCG	CTC	CTG	TCC	ACA	GGA	ATC	ATC	TGC	ACC	TTG	TAC	ACT	646
Ala	Ser	Leu	Leu	Ser	Thr	Gly	Ile	Ile	Cys	Thr	Leu	Tyr	Thr	
				170					175					
ACC	GTG	GGT	GGT	ATG	AAG	GCC	GTG	GTC	TGG	ACA	GAT	GTG	TTC	688
Thr	Val	Gly	Gly	Met	Lys	Ala	Val	Val	Trp	Thr	Asp	Val	Phe	
180					185					190				
CAG	GTT	GTG	GTA	ATG	CTC	GTT	GGC	TTC	TGG	GTG	ATC	CTG	GCC	730
Gln	Val	Val	Val	Met	Leu	Val	Gly	Phe	Trp	Val	Ile	Leu	Ala	
	195					200					205			
CGA	GGC	GTC	ATT	CTC	CTG	GGG	GGT	CCC	CGG	AAC	GTG	CTC	AGC	772
Arg	Gly	Val	Ile	Leu	Leu	Gly	Gly	Pro	Arg	Asn	Val	Leu	Ser	
		210					215					220		
CTC	GCT	CAG	AAC	CAT	TCC	CGG	ATC	AAC	CTG	ATG	GAC	TTT	GAC	814
Leu	Ala	Gln	Asn	His	Ser	Arg	Ile	Asn	Leu	Met	Asp	Phe	Asp	
			225					230					235	
CCT	GAT	CCT	CGG	AGC	CGG	TAC	ACC	TTC	TGG	ACT	TTC	ATA	GTG	856
Pro	Asp	Pro	Arg	Ser	Arg	Tyr	Thr	Phe	Trp	Thr	Phe	Ile	Val	
				240					245					
GGT	GGC	ACA	CTG	GTG	TGG	CTC	TCC	ATG	TAC	GGT	GTG	AAC	CAA	898
Gly	Gly	Thr	Leu	Val	Trp	Leu	Ser	Met	Tyr	Gly	Val	Asn	Gln	
250					255					260				
GCC	CAG	GTA	CAG	CGC	TAT	GTG	GCC	TGC	CAC	ACA	GAG	GGA	AAG	940
Ala	Gln	Val	Gln	Arg	Tyr	Val	Ala	Cys	His	Thr	Glu	Gly	Lys	
	265					270					275			
GCC	AAA	CTG	GCC	CTG	CTT	GTC	AAC	CAG	CTG	GGC	CTC	TTC	CTG	982
Ala	Lys	Leu	Ala	Leu	Leu	Val	Asn	Gln	Leu	Gly	Leu	Phe	Leu	
		280					285					290		
ATT	GTG	GCC	AGT	GCA	GCT	TGC	TGT	GGC	ATT	GTC	ATG	TTC	GTC	1024
Ile	Val	Ala	Ser	Ala	Ala	Cys	Cys	Gly	Ile	Val	Met	Phe	Val	
			295					300					305	

TAC	TAC	AAG	GAC	TGT	GAC	CCC	CTC	CTC	ACA	GGC	CGT	ATC	TCA	1066
Tyr	Tyr	Lys	Asp	Cys	Asp	Pro	Leu	Leu	Thr	Gly	Arg	Ile	Ser	
			310						315					
GCC	CCC	GAC	CAG	TAC	ATG	CCG	CTG	CTT	GTG	TTG	GAC	ATT	TTT	1108
Ala	Pro	Asp	Gln	Tyr	Met	Pro	Leu	Leu	Val	Leu	Asp	Ile	Phe	
320					325					330				
GAG	GAT	CTG	CCC	GGG	GTC	CCC	GGG	CTC	TTC	CTG	GCC	TGT	GCC	1150
Glu	Asp	Leu	Pro	Gly	Val	Pro	Gly	Leu	Phe	Leu	Ala	Cys	Ala	
	335					340					345			
TAC	AGT	GGC	ACC	CTC	AGC	ACT	GCA	TCC	ACC	AGC	ATC	AAC	GCC	1192
Tyr	Ser	Gly	Thr	Leu	Ser	Thr	Ala	Ser	Thr	Ser	Ile	Asn	Ala	
		350					355					360		
ATG	GCA	GCT	GTG	ACT	GTG	GAA	GAC	CTC	ATC	AAG	CCG	AGG	ATG	1234
Met	Ala	Ala	Val	Thr	Val	Glu	Asp	Leu	Ile	Lys	Pro	Arg	Met	
			365					370					375	
CCT	GGC	CTG	GCA	CCT	CGG	AAG	TTG	GTT	TTC	ATC	TCT	AAA	GGG	1276
Pro	Gly	Leu	Ala	Pro	Arg	Lys	Leu	Val	Phe	Ile	Ser	Lys	Gly	
				380					385					
CTC	TCA	TTC	ATC	TAC	GGC	TCT	GCC	TGC	CTC	ACT	GTG	GCT	GCT	1318
Leu	Ser	Phe	Ile	Tyr	Gly	Ser	Ala	Cys	Leu	Thr	Val	Ala	Ala	
390					395					400				
CTG	TCC	TCA	CTG	CTG	GGA	GGT	GGT	GTC	CTC	CAG	GGT	TCC	TTC	1360
Leu	Ser	Ser	Leu	Leu	Gly	Gly	Gly	Val	Leu	Gln	Gly	Ser	Phe	
	405					410					415			
ACT	GTG	ATG	GGT	GTC	ATC	AGT	GGG	CCT	CTA	CTA	GGC	GCC	TTC	1402
Thr	Val	Met	Gly	Val	Ile	Ser	Gly	Pro	Leu	Leu	Gly	Ala	Phe	
		420					425					430		
ACG	CTT	GGG	ATG	CTG	CTC	CCA	GCC	TGC	AAC	ACG	CCA	GGC	GTT	1444
Thr	Leu	Gly	Met	Leu	Leu	Pro	Ala	Cys	Asn	Thr	Pro	Gly	Val	
			435					440					445	
CTC	TCC	GGG	TTG	GCA	GCA	GGC	TTG	GCT	GTA	TCC	CTG	TGG	GTG	1486
Leu	Ser	Gly	Leu	Ala	Ala	Gly	Leu	Ala	Val	Ser	Leu	Trp	Val	
				450					455					
GCC	GTA	GGG	GCC	ACA	CTG	TAT	CCC	CCT	GGA	GAG	CAG	ACC	ATG	1528
Ala	Val	Gly	Ala	Thr	Leu	Tyr	Pro	Pro	Gly	Glu	Gln	Thr	Met	
460					465					470				
GGG	GTG	CTG	CCC	ACC	TCG	GCT	GCA	GGC	TGC	ACC	AAC	GAT	TCG	1570
Gly	Val	Leu	Pro	Thr	Ser	Ala	Ala	Gly	Cys	Thr	Asn	Asp	Ser	
	475					480					485			



AGGAGACAGT AGAAGGGTCA TAGATACAAA GAAAACTAAG GCAGAGGGAG	2343
AAATGAATTG TCTACAGAGC ACAGAGCTCC AAGGATTGTG AAGCTACCTT	2393
GAGGTGCCAA GGGACGGATT CTCAGAGCCT TCACAAGACA CAAACGGACG	2443
AGTTGCCTCC TCCAATTCAG ATGGTTTGCA GACTATCAGA GAACATGTTT	2493
CTCCTGTGAT CAGCTACCTA GCCTCTGCCA ACGTGTTCCA GCTTCCAGGA	2543
GGCCACACAG ACCCCACCCC CCATGCTCTC ACCCTTTACC CCTGTGCTTT	2593
TCACACACTA GGCAACTGCT CCACCACAGG ACCTCACACC TAGACCTCCG	2643
TTTTTGACAC AGGGCCTTAA GGTAATCTGG CTGCCATCTG ACTATCTCTC	2693
AGCACGTTCA CGTGTACAAT ATTTCAATTCT TTTTCATTGC CAAGTTGTCT	2743
TGTAAGGAGA GACCACAATG TGTCATCCAT GCCCAGCTTT TGTGTCTAAC	2793
AAATAAAATC GCTGAAGGTG TTCAGGTGCA ATGGCCTGTG ACATTA	2839

(3) INFORMATION FOR SEQ ID NO: 2

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 618
  - (B) TYPE: AMINO ACID
  - (D) TOPOLOGY: LINEAR
- (ii) MOLECULE TYPE:
  - (A) DESCRIPTION: PROTEIN
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: RAT
  - (B) INDIVIDUAL ISOLATE: SODIUM/IODIDE SYMPORTER
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2

Met	Glu	Gly	Ala	Glu	Ala	Gly	Ala	Arg	Ala	Thr
1				5					10	
Phe	Gly	Ala	Trp	Asp	Tyr	Gly	Val	Phe	Ala	Thr
15								20		25
Val	Ser	Thr	Gly	Ile	Gly	Leu	Trp	Val	Gly	Leu
			30					35		

Gly	Gln	Arg	Ser	Ala	Asp	Asp	Phe	Phe	Thr	Gly	Gly	Arg	Gln
40					45					50			
Leu	Ala	Ala	Val	Pro	Val	Gly	Leu	Ser	Leu	Ala	Ala	Ser	Phe
	55					60					65		
Met	Ser	Ala	Val	Gln	Val	Leu	Gly	Val	Pro	Ala	Glu	Ala	Ala
		70					75					80	
Arg	Tyr	Gly	Leu	Lys	Phe	Leu	Trp	Met	Cys	Ala	Gly	Gln	Leu
			85					90					95
Leu	Asn	Ser	Leu	Leu	Thr	Ala	Phe	Leu	Phe	Leu	Pro	Ile	Phe
				100					105				
Tyr	Arg	Leu	Gly	Leu	Thr	Ser	Thr	Tyr	Gln	Tyr	Leu	Glu	Leu
110					115					120			
Arg	Phe	Ser	Arg	Ala	Val	Arg	Leu	Cys	Gly	Thr	Leu	Gln	Tyr
	125					130					135		
Leu	Val	Ala	Thr	Met	Leu	Tyr	Thr	Gly	Ile	Val	Ile	Tyr	Ala
		140					145					150	
Pro	Ala	Leu	Ile	Leu	Asn	Gln	Val	Thr	Gly	Leu	Asp	Ile	Trp
			155					160					165
Ala	Ser	Leu	Leu	Ser	Thr	Gly	Ile	Ile	Cys	Thr	Leu	Tyr	Thr
				170					175				
Thr	Val	Gly	Gly	Met	Lys	Ala	Val	Val	Trp	Thr	Asp	Val	Phe
180					185					190			
Gln	Val	Val	Val	Met	Leu	Val	Gly	Phe	Trp	Val	Ile	Leu	Ala
	195					200					205		
Arg	Gly	Val	Ile	Leu	Leu	Gly	Gly	Pro	Arg	Asn	Val	Leu	Ser
		210					215					220	
Leu	Ala	Gln	Asn	His	Ser	Arg	Ile	Asn	Leu	Met	Asp	Phe	Asp
			225					230					235
Pro	Asp	Pro	Arg	Ser	Arg	Tyr	Thr	Phe	Trp	Thr	Phe	Ile	Val
				240					245				
Gly	Gly	Thr	Leu	Val	Trp	Leu	Ser	Met	Tyr	Gly	Val	Asn	Gln
250					255					260			
Ala	Gln	Val	Gln	Arg	Tyr	Val	Ala	Cys	His	Thr	Glu	Gly	Lys
	265					270					275		

Ala	Lys	Leu	Ala	Leu	Leu	Val	Asn	Gln	Leu	Gly	Leu	Phe	Leu
		280					285					290	
Ile	Val	Ala	Ser	Ala	Ala	Cys	Cys	Gly	Ile	Val	Met	Phe	Val
			295					300					305
Tyr	Tyr	Lys	Asp	Cys	Asp	Pro	Leu	Leu	Thr	Gly	Arg	Ile	Ser
				310					315				
Ala	Pro	Asp	Gln	Tyr	Met	Pro	Leu	Leu	Val	Leu	Asp	Ile	Phe
320					325					330			
Glu	Asp	Leu	Pro	Gly	Val	Pro	Gly	Leu	Phe	Leu	Ala	Cys	Ala
	335					340					345		
Tyr	Ser	Gly	Thr	Leu	Ser	Thr	Ala	Ser	Thr	Ser	Ile	Asn	Ala
		350					355					360	
Met	Ala	Ala	Val	Thr	Val	Glu	Asp	Leu	Ile	Lys	Pro	Arg	Met
			365					370					375
Pro	Gly	Leu	Ala	Pro	Arg	Lys	Leu	Val	Phe	Ile	Ser	Lys	Gly
				380					385				
Leu	Ser	Phe	Ile	Tyr	Gly	Ser	Ala	Cys	Leu	Thr	Val	Ala	Ala
390					395					400			
Leu	Ser	Ser	Leu	Leu	Gly	Gly	Gly	Val	Leu	Gln	Gly	Ser	Phe
	405					410					415		
Thr	Val	Met	Gly	Val	Ile	Ser	Gly	Pro	Leu	Leu	Gly	Ala	Phe
		420					425					430	
Thr	Leu	Gly	Met	Leu	Leu	Pro	Ala	Cys	Asn	Thr	Pro	Gly	Val
			435					440					445
Leu	Ser	Gly	Leu	Ala	Ala	Gly	Leu	Ala	Val	Ser	Leu	Trp	Val
				450					455				
Ala	Val	Gly	Ala	Thr	Leu	Tyr	Pro	Pro	Gly	Glu	Gln	Thr	Met
460					465					470			
Gly	Val	Leu	Pro	Thr	Ser	Ala	Ala	Gly	Cys	Thr	Asn	Asp	Ser
	475					480					485		
Val	Leu	Leu	Gly	Pro	Pro	Gly	Ala	Thr	Asn	Ala	Ser	Asn	Gly
		490					495					500	
Ile	Pro	Ser	Ser	Gly	Met	Asp	Thr	Gly	Arg	Pro	Ala	Leu	Ala
			505					510					515



Asp Thr Phe Tyr Ala Ile Ser Tyr Leu Tyr Tyr Gly Ala Leu  
520 525

Gly Thr Leu Thr Thr Met Leu Cys Gly Ala Leu Ile Ser Tyr  
530 535 540

Leu Thr Gly Pro Thr Lys Arg Ser Ser Leu Gly Pro Gly Leu  
545 550 555

Leu Trp Trp Asp Leu Ala Arg Gln Thr Ala Ser Val Ala Pro  
560 565 570

Lys Glu Asp Thr Ala Thr Leu Glu Glu Ser Leu Val Lys Gly  
575 580 585

Pro Glu Asp Ile Pro Ala Val Thr Lys Lys Pro Pro Gly Leu  
590 595

Lys Pro Gly Ala Glu Thr His Pro Leu Tyr Leu Gly His Asp  
600 605 610

Val Glu Thr Asn Leu  
615